AGGCTGTCCCACCCACCATCTGCACCCGCTGCAGCGCCCGCGCCCCTGTCCCGCGCCGTAGTCGTCATTTGTAGCCC
GCCTGCCGCTCCCGGGGACCCGATCCTACCCTGGGTGCGGGGCAGAGCGGGCATGGCCCGTCTGGGGACCGCCTGCC
CTGCGCTGGCGCTGGCCCTGGCACTTGTGGCGGTGGCCCTGGCTGGAGTCAGAGCCCAGGGCGCAGCCTTCGAGGAG
CCTGACTATTACAGCCAGGAGCTCTGGCGGCGCGGGCGCTATTATGGGCATCCGGAGCCTGAGCCGGAGCCGGAGCT
CTTCTCGCCTTCAATGCATGAAGACCTTAGGGTGGAGGAGCAGGAACAGCAGGAGCCGCACCAGCAGGGCCACAGAA
CTCCCAAGAAGGCCATCAAGCCCAAGAAGGCTCCCAAGAGGGAGAAGTTAGTTGCAGAGACGCCTCCACCAGGTAAA
AATAGCAACAGAAAAGGCAGAAGAAGCAAGAATCTTGAGAAAGCTGCCAGTGATGACCATGGTGTCCCTGTGGCTCA
TGAGGATGTCAGAGAGAGTTGCCCACCTCTTGGTCTGGAAACATTAAAAATCACAGACTTCCAGCTGCATGCCTCCA
CATCGAAGCGTTATGGCCTGGGAGCCCACCGGGGGAGACTCAACATCCAGGCAGGCATTAATGAAAATGACTTTTAC
GATGGGGCTTGGTGTGCTGGTAGGAACGACTTGCATCAGTGGATCGAAGTGGATGCCCGGCGCCTGACCAAGTTCAC
AGGGGTCATTACCCAAGGAAGGAACTCTCTCTGGCTGAGTGACTGGGTGACATCCTATAAAGTCATGGTGAGCAATG
ACAGCCACACATGGGTTACTGTGAAGAATGGATCTGGCGACATGATATTTGAAGGAAACAGTGAGAAGGAGATTCCT
GTGCTCAATGAGCTGCCAGTCCCCATGGTGGCCCGCTACATTCGCATAAACCCTCAGTCCTGGTTTGATAACGGGAG
CATCTGCATGAGGATGGAGATCTTGGGCTGCCCACTGCCGGATCCTAATAACTATTATCACCGACGTAATGAGATGA
CCACCACGGATGACCTGGATTTTAAGCACCACAACTATAAGGAAATGCGCCAGTTGATGAAGGTTGTCAATGAAATG
TGCCCCAATATTACCAGGATTTACAACATTGGCAAAAGCCACCAGGGCCTGAAATTGTATGCGGTAGAGATCTCTGA
CCATCCTGGGGAACATGAAGTTGGTGAGCCCGAGTTCCACTACATCGCAGGGGCCCACGGCAATGAGGTTCTGGGAC
GAGAACTGCTGCTGCTGCTGCACTTCCTCTGCCAGGAATACTCGGCGCAGAACGCACGCATCGTCCGCTTGGTG
GAGGAGACTCGAATCCACATTCTACCCTCCCTCAATCCTGATGGCTATGAGAAGGCCTATGAAGGAGGTTCCGAGTT
GGGAGGCTGGTCCCTGGGACGTTGGACCCATGATGGCATCGATATCAACAACAACTTTCCGGATTTAAACTCGCTGC
TCTGGGAGGCAGAGGACCAGCAGAATGCCCCAAGGAAGGTCCCCAACCACTACATTGCCATCCCTGAGTGGTTTCTG
TCTGAGAATGCCACAGTGGCCACAGAGACCAGAGCCGTCATCGCCTGGATGGAGAAGATCCCGTTTGTGCTGGGAGG
CAACCTACAGGGGGGTGAGCTGGTCGTGGCATACCCCTATGACATGGTGCGGTCCCTGTGGAAGACCCAGGAGCACA
CCCCAACACCTGATGATCATGTGTTCCGCTGGCTGGCGTATTCCTACGCCTCCACTCACCGCCTCATGACAGATGCC
AGGAGGCGAGTGTGCCACACGGAAGATTTTCAGAAGGAGGAGGGCACCGTCAATGGGGCTTCCTGGCACACAGTGGC
TGGAAGTCTAAACGATTTCAGCTACCTCCATACAAACTGCTTTGAGCTGTCCATCTACGTGGGCTGTGATAAATACC
CACACGAGAGCGAGCTGCCGGAGGAATGGGAGAATAACCGGGAGTCTCTGATTGTGTTCATGGAGCAGGTTCATCGA
GGCATCAAAGGCATAGTGAGAGATTTACAAGGGAAAGGGATTTCAAATGCTGTCATCTCTGTGGAAGGTGTTAACCA
TGACATCCGGACAGCCAGCGATGGGGATTACTGGCGTCTACTGAACCCTGGCGAATATGTGGTCACAGCCAAGGCGG
AAGGCTTTATCACTTCCACCAAGAACTGCATGGTTGGCTATGATATGGGAGCTACTCGGTGTGACTTCACCCTCACA
AAGACCAACCTGGCTAGGATAAGAGAAATTATGGAGACATTTGGGAAGCAGCCTGTCAGCCTACCCTCCAGGCGCCT
GAAGCTGCGGGGACGGAAAAGGCGGCAGCGTGGGTGACCCTGTCGGACACTTGAGACATACCCCAGACCGTGCAAAT
AAAAATCCACTCCAGTAGTAAAAAAA
(SEQ ID NO:1)

MARLGTACPALALALALVAVALAGVRAQGAAFEEPDYYSQELWRRGRYYGHPEPEPEPELFSPSMHEDLR
VEEQEQQEPHQQGHRTPKKAIKPKKAPKREKLVAETPPPGKNSNRKGRRSKNLEKAASDDHGVPVAHEDV
RESCPPLGLETLKITDFQLHASTSKRYGLGAHRGRLNIQAGINENDFYDGAWCAGRNDLHQWIEVDARRL
TKFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIFEGNSEKEIPVLNELPVPMVARYIR
INPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTTTDDLDFKHHNYKEMRQLMKVVNEMCPNITRIY
NIGKSHQGLKLYAVEISDHPGEHEVGEPEFHYIAGAHGNEVLGRELLLLLLHFLCQEYSAQNARIVRLVE
ETRIHILPSLNPDGYEKAYEGGSELGGWSLGRWTHDGIDINNNFPDLNSLLWEAEDQQNAPRKVPNHYIA
IPEWFLSENATVATETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDMVRSLWKTQEHTPTPDDHVFRWLA
YSYASTHRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHESEL
PEEWENNRESLIVFMEQVHRGIKGIVRDLQGKGISNAVISVEGVNHDIRTASDGDYWRLLNPGEYVVTAK
AEGFITSTKNCMVGYDMGATRCDFTLTKTNLARIREIMETFGKQPVSLPSRRLKLRGRKRRQRG (SEQ
ID NO:2)

Figure 1

underlined = deleted in targeting construct

[ ] = sequence flanking Neo insert in targeting construct

```
[AGGCTGTCCCACCCACCATCTGCACCCGCTGCAGCGCCCGCGCCCCTGTCCCGCGCCGT
AGTCGTCATTTGTAGCCCGCCTGCCGCTCCCGGGGACCCGATCCTACCCTGGGTGCGGGG
CAGAGCGGGCATGGCCCGTCTGGGGACCGCCTGCCCTGCGCTGGCGCTGGCCCTGGCACT
TGTGGCGGTGGCCCTGGCTGGAGTCAGAGCCCAGGGCGCAGCCTTCGAGGAGCCTGACTA
TTACAGCCAGGAGCTCTGGCGGCGCGGGCGCTATTATGGGCATCCGGAGCCTGAGCCGGA
GCCGGAGCTCTTCTCGCCTTCAATGCA]TGAAGACCTTAGGGTGGAGGAGCAGGAACAGC
AGGAGCCGCACCAGCAGGGCCACAGAACTCCCAAGAAGGCCATCAAGCCCAAGAAGGCTC
CCAA[GAGGGAGAAGTTAGTTGCAGAGACGCCTCCACCAG]GTAAAAATAGCAACAGAAA
AGGCAGAAGAAGCAAGAATCTTGAGAAAGCTGCCAGTGATGACCATGGTGTCCCTGTGGC
TCATGAGGATGTCAGAGAGAGTTGCCCACCTCTTGGTCTGGAAACATTAAAAATCACAGA
CTTCCAGCTGCATGCCTCCACATCGAAGCGTTATGGCCTGGGAGCCCACCGGGGGGAGACT
CAACATCCAGGCAGGCATTAATGAAAATGACTTTTACGATGGGGCTTGGTGTGCTGGTAG
GAACGACTTGCATCAGTGGATCGAAGTGGATGCCCGGCGCCTGACCAAGTTCACAGGGGT
CATTACCCAAGGAAGGAACTCTCTCTGGCTGAGTGACTGGGTGACATCCTATAAAGTCAT
GGTGAGCAATGACAGCCACACATGGGTTACTGTGAAGAATGGATCTGGCGACATGATATT
TGAAGGAAACAGTGAGAAGGAGATTCCTGTGCTCAATGAGCTGCCAGTCCCCATGGTGGC
CCGCTACATTCGCATAAACCCTCAGTCCTGGTTTGATAACGGGACGCATCTGCATGAGGAT
GGAGATCTTGGGCTGCCCACTGCCGGATCCTAATAACTATTATCACCGACGTAATGAGAT
GACCACCACGGATGACCTGGATTTTAAGCACCACAACTATAAGGAAATGCGCCAGTTGAT
GAAGGTTGTCAATGAAATGTGCCCCAATATTACCAGGATTTACAACATTGGCAAAAGCCA
CCAGGGCCTGAAATTGTATGCGGTAGAGATCTCTGACCATCCTGGGGAACATGAAGTTGG
TGAGCCCGAGTTCCACTACATCGCAGGGGCCCACGGCAATGAGGTTCTGGGACGAGAACT
GCTGCTGCTGCTGCTGCACTTCCTCTGCCAGGAATACTCGGCGCAGAACGCACGCATCGT
CCGCTTGGTGGAGGAGACTCGAATCCACATTCTACCCTCCCTCAATCCTGATGGCTATGA
GAAGGCCTATGAAGGAGGTTCCGAGTTGGGAGGCTGGTCCCTGGGACGTTGGACCCATGA
TGGCATCGATATCAACAACAACTTTCCGGATTTAAACTCGCTGCTCTGGGAGGCAGAGGA
CCAGCAGAATGCCCCAAGGAAGGTCCCCAACCACTACATTGCCATCCCTGAGTGGTTTCT
GTCTGAGAATGCCACAGTGGCCACAGAGACCAGAGCCGTCATCGCCTGGATGGAGAAGAT
CCCGTTTGTGCTGGGAGGCAACCTACAGGGGGGTGAGCTGGTCGTGGCATACCCCTATGA
CATGGTGCGGTCCCTGTGGAAGACCCAGGAGCACACCCCAACACCTGATGATCATGTGTT
CCGCTGGCTGGCGTATTCCTACGCCTCCACTCACCGCCTCATGACAGATGCCAGGAGGCG
AGTGTGCCACACGGAAGATTTTCAGAAGGAGGAGGGCACCGTCAATGGGGCTTCCTGGCA
CACAGTGGCTGGAAGTCTAAACGATTTCAGCTACCTCCATACAAACTGCTTTGAGCTGTC
CATCTACGTGGGCTGTGATAAATACCCACACGAGAGCGAGCTGCCGGAGGAATGGGAGAA
TAACCGGGAGTCTCTGATTGTGTTCATGGAGCAGGTTCATCGAGGCATCAAAGGCATAGT
GAGAGATTTACAAGGGAAAGGGATTTCAAATGCTGTCATCTCTGTGGAAGGTGTTAACCA
TGACATCCGGACAGCCAGCGATGGGGATTACTGGCGTCTACTGAACCCTGGCGAATATGT
GGTCACAGCCAAGGCGGAAGGCTTTATCACTTCCACCAAGAACTGCATGGTTGGCTATGA
TATGGGAGCTACTCGGTGTGACTTCACCCTCACAAAGACCAACCTGGCTAGGATAAGAGA
AATTATGGAGACATTTGGGAAGCAGCCTGTCAGCCTACCCTCCAGGCGCCTGAAGCTGCG
GGGACGGAAAAGGCGGCAGCGTGGGTGACCCTGTCGGACACTTGAGACATACCCCAGACC
GTGCAAATAAAAATCCACTCCAGTAGTAAAAAA (SEQ ID NO:1)
```
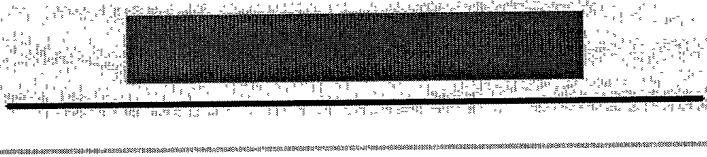
Figure 2A

Gene Sequence Structure *

327 bp   Sequence Deleted   422 bp

Size of full-length cDNA: 2490 bp
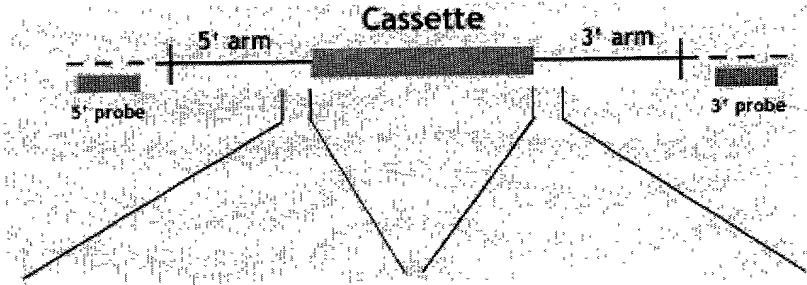

Targeting Vector* (genomic sequence)

LacZ-Neo

Construct Number: 2825

Arm Length:
5': 4 kb
3': 1.5 kb

──────── Targeting Vector
─ ─ ─ ─ ─ Endogenous Locus
* Not drawn to scale

5'>GGCATGGCCCGTCTGGGGACC
GCCTGCCCTGCGCTGGCGCTGGCC
CTGGCACTTGTGGCGGTGGCCCTG
GCTGGAGTCAGAGCCCAGGGCGCA
GCCTTCGAGGAGCCTGACTATTAC
AGCCAGGAGCTCTGGCGGCGCGGG
CGCTATTATGGGCATCCGGAGCCT
GAGCCGGAGCCGGAGCTCTTCTCG
CCTTCAATGCA<3'
(SEQ ID NO:3)

5'>GAGGGAGAAGTTAGTTGCAGA
GACGCCTCCACCAGGTAACTTTTG
CATCGGGCAGCCCGAGGGGGCGCC
AGCGATCGTGGCACTCCAGGGGAC
ACCTGGCTTCCAGTATGTTTTCTT
GAGTGAGCCCAGCCAAAGTCCTGT
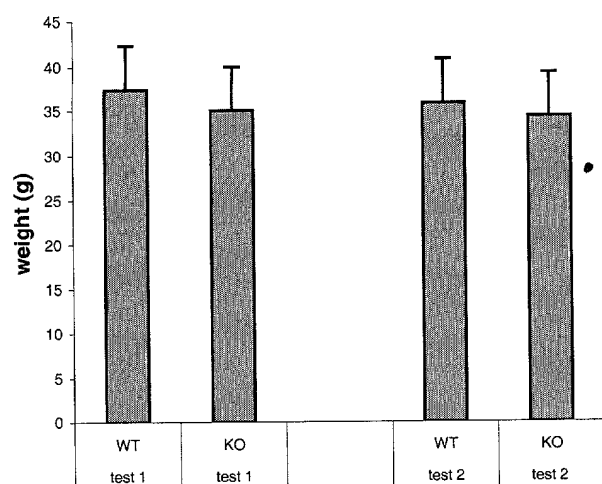GGTGCCTGTGTTATTCCCTAGAGA
CTACATCTGAGCTAAGTTCAGCTT
TCTCTCCCTGC<3'
(SEQ ID NO:4)

Figure 2B
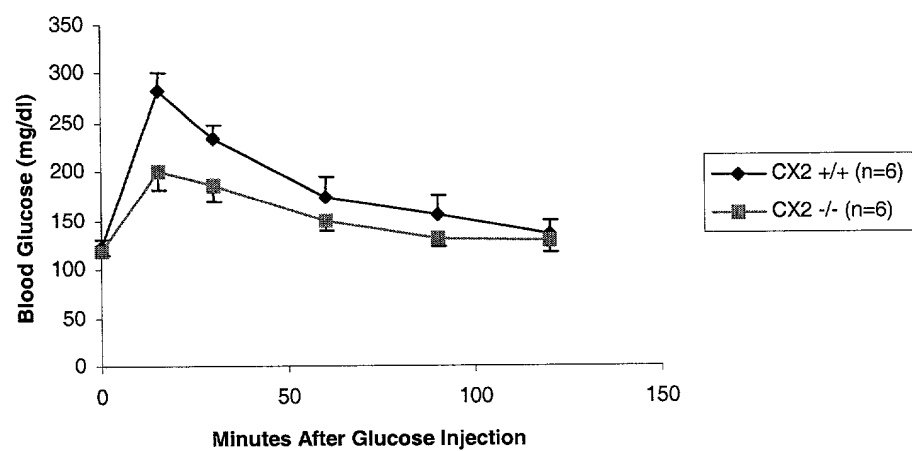
Figure 3

Figure 4